

ACCESSION NUMBER: 1999:173061 CAPLUS  
TITLE: Influence of Centriole Behavior on the First Spindle  
Formation in Zygotes of the Brown Alga *Fucus distichus*  
(Fucales, Phaeophyceae)  
AUTHOR(S): Nagasato, Chikako; Motomura, Taizo; Ichimura, Terunobu  
CORPORATE SOURCE: Institute of Algological Research, Faculty of Science,  
Hokkaido University, Muroran, 051-0003, Japan  
SOURCE: Developmental Biology (1999), 208(1), 200-209  
CODEN: DEBIAO; ISSN: 0012-1606  
PUBLISHER: Academic Press  
DOCUMENT TYPE: Journal  
LANGUAGE: English  
REFERENCE COUNT: 54 THERE ARE 54 CITED REFERENCES AVAILABLE FOR THIS  
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 22, 2005, 23:23:20 ; Search time 217 Seconds  
(without alignments)  
3316.871 Million cell updates/sec

Title: US-09-914-698-1  
Perfect score: 9514  
Sequence: 1 MELVWSPVLEVACKETLQLI.....FISSVYAFDTILCKLQIDMF 1861

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	9514	100.0	1861	3	AAY90350	Aay90350 Drosophil
2	9482	99.7	1954	4	ABB62757	Abb62757 Drosophil
3	755	7.9	1277	8	ADQ65753	Adq65753 Novel hum
4	672.5	7.1	898	7	ADM04714	Adm04714 Human pro
5	541	5.7	934	5	ABP64732	Abp64732 Human pro
6	408.5	4.3	726	4	AAB92933	Aab92933 Human pro
7	408.5	4.3	726	6	ABR48183	Abr48183 Human bla
8	408.5	4.3	726	8	ADN04938	Adn04938 Antipsori
9	319.5	3.4	526	4	AAB92906	Aab92906 Human pro

10	303.5	3.2	2779	4	ABB62371	Abb62371	Drosophil
11	302.5	3.2	4576	8	ADO55184	Ado55184	Protein #
12	300.5	3.2	1855	6	ABR47539	Abr47539	Breast ca
13	300.5	3.2	1855	8	ADN04056	Adn04056	Antipsori
14	294	3.1	4684	8	ADH09904	Adh09904	Human hos
15	294	3.1	4684	8	ADH09906	Adh09906	Human hos
16	291	3.1	1742	7	ADB82767	Adb82767	Human pro
17	290.5	3.1	285	4	AAB92905	Aab92905	Human pro
18	290.5	3.1	285	8	ADQ18923	Adq18923	Human sof
19	284.5	3.0	1805	6	ABU35589	Abu35589	Protein e
20	284.5	3.0	1805	7	ABO23572	Abo23572	Mycoplasma
21	282.5	3.0	1584	8	ABM80133	Abm80133	Tumour-as
22	277	2.9	1818	6	ABU36321	Abu36321	Protein e
23	277	2.9	1898	2	AAY30795	Aay30795	A human t
24	277	2.9	1898	7	ADD48869	Add48869	Human Pro
25	275.5	2.9	2688	4	AAM40883	Aam40883	Human pol
26	271.5	2.9	1975	4	ABB62094	Abb62094	Drosophil
27	270	2.8	2383	5	ABG65631	Abg65631	Human bre
28	269.5	2.8	2663	4	AAM39097	Aam39097	Human pol
29	269	2.8	1525	6	AAE32042	Aae32042	Human ves
30	268	2.8	2482	2	AAR72826	Aar72826	Human mit
31	268	2.8	2482	2	AAW23996	Aaw23996	Human mit
32	268	2.8	3113	7	ADN95402	Adn95402	Human BEC
33	268	2.8	3113	8	ADK70220	Adk70220	Human oes
34	268	2.8	3113	8	ADL72172	Adl72172	Human sol
35	268	2.8	3113	8	ADQ18045	Adq18045	Human sof
36	267.5	2.8	2663	8	ADQ17932	Adq17932	Human sof
37	265.5	2.8	2633	4	ABG06505	Abg06505	Novel hum
38	264	2.8	1530	8	ADL13007	Adl13007	Human ste
39	264	2.8	2442	3	AAY77575	Aay77575	Human cyt
40	260	2.7	1979	7	ADB75595	Adb75595	Prostate
41	259	2.7	3248	2	AAR99795	Aar99795	Kinetocho
42	257.5	2.7	1374	4	AAB69070	Aab69070	Human mal
43	257.5	2.7	1854	8	ADR66477	Adr66477	Human pro
44	257.5	2.7	1854	8	ADR66135	Adr66135	Human pro
45	255.5	2.7	2017	4	ABG06301	Abg06301	Novel hum

# ALIGNMENTS

RESULT 1  
 AAY90350  
 ID AAY90350 standard; protein; 1861 AA.  
 XX  
 AC AAY90350;  
 XX  
 DT 04-DEC-2000 (first entry)  
 XX  
 DE Drosophila Asp protein sequence.  
 XX  
 KW Asp; Drosophila; microtubule organising centre; MTOC; mitosis inhibitor;  
 KW tumour cell.  
 XX  
 OS Drosophila sp.  
 XX  
 PN WO200052478-A1.

XX  
 PD 08-SEP-2000.  
 XX  
 PF 03-MAR-2000; 2000WO-GB000785.  
 XX  
 PR 04-MAR-1999; 99GB-00005007.  
 XX  
 PA (UYDU-) UNIV DUNDEE.  
 XX  
 PI Glover DM, Avides MDC;  
 XX  
 DR WPI; 2000-594203/56.  
 DR N-PSDB; AAA37761.  
 XX  
 PT Use of Drosophila Asp polypeptide for identifying substances capable of  
 PT disrupting microtubule organizing center integrity and use of the  
 PT identified substances for inhibiting mitosis in tumor cell.  
 XX  
 PS Claim 4; Page 43-44; 51pp; English.  
 XX  
 CC This sequence represents the Drosophila Asp protein. The invention  
 CC relates to the use of Drosophila Asp polypeptide (or its homologue, or  
 CC fragment) capable of stimulating formation and/or maintenance of  
 CC microtubule organising centres (MTOCs), in an assay for identifying a  
 CC substance capable of disrupting MTOC integrity. Asp polypeptide or its  
 CC homolog is useful for identifying a substance capable of disrupting MTOC  
 CC integrity. Substances identified by the method can be used to inhibit  
 CC mitosis, e.g. in tumour cells  
 XX  
 SQ Sequence 1861 AA;

Query Match 100.0%; Score 9514; DB 3; Length 1861;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MELVWSPVLEVACKETLQLIDNRNFRKEVMIILKSKSNQPVKNPRKFPTVGKTLQLKSPT	60
Db	1	MELVWSPVLEVACKETLQLIDNRNFRKEVMIILKSKSNQPVKNPRKFPTVGKTLQLKSPT	60
Qy	61	GAGKTMKSVVSAAVQQKKRMSAAAAPPSKQTRVTAPSRPAAWAHPPPQAPLVEKNVYKT	120
Db	61	GAGKTMKSVVSAAVQQKKRMSAAAAPPSKQTRVTAPSRPAAWAHPPPQAPLVEKNVYKT	120
Qy	121	PQEEPVIYISPPRSLKENLSPMTPGNLLDVIDNLRFTPLTETRKGQATIFPDNLAAWPT	180
Db	121	PQEEPVIYISPPRSLKENLSPMTPGNLLDVIDNLRFTPLTETRKGQATIFPDNLAAWPT	180
Qy	181	PTLKGNVKSCANDMRPRRITPDDLEDQPATNKTFDVKHSETINISLDTLDCSRIDGQPHT	240
Db	181	PTLKGNVKSCANDMRPRRITPDDLEDQPATNKTFDVKHSETINISLDTLDCSRIDGQPHT	240
Qy	241	PLNKTTTIVHATHTRALACIHEEEGSPSPRTPTKSAIHDLKRDIKLVGSPLRKYESMKD	300
Db	241	PLNKTTTIVHATHTRALACIHEEEGSPSPRTPTKSAIHDLKRDIKLVGSPLRKYESMKD	300
Qy	301	LSLLSPQTKYAIQGSMPNLNEMKIRSIEQNRYQQEQIQIKAKDLNSSSSSEASLAGQQE	360

Db	301	LSLLSPQTKYAIQGSMPNLNEMKIRSIEQNRYQEQQIQIKAKDLNSSSSSEASLAGQQE	360
Qy	361	FLFNHSEILAQSSRNLHEVGRKSVKGSVKPNPHKRRSHELFSFDAPSNESELYRNETVAI	420
Db	361	FLFNHSEILAQSSRNLHEVGRKSVKGSVKPNPHKRRSHELFSFDAPSNESELYRNETVAI	420
Qy	421	SPPKKQRVEDTTLPRSAAPANASARSSSAHAWPHAQSKKFKLAQTMSLMKKPATPRKVRD	480
Db	421	SPPKKQRVEDTTLPRSAAPANASARSSSAHAWPHAQSKKFKLAQTMSLMKKPATPRKVRD	480
Qy	481	TSIQPSVKLYDSELYMQTCINPDPFAATTTIDPFLASTMYLDEQAVDRHQADFCKKWLNAL	540
Db	481	TSIQPSVKLYDSELYMQTCINPDPFAATTTIDPFLASTMYLDEQAVDRHQADFCKKWLNAL	540
Qy	541	VSIPADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNYLTKYRLETLRKAAVELFF	600
Db	541	VSIPADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNYLTKYRLETLRKAAVELFF	600
Qy	601	SEQMRLPCSKVAVYVVKQALRIRSDRNHLHDVVMQRTILELLLCFNPLWLRLGLEVVFE	660
Db	601	SEQMRLPCSKVAVYVVKQALRIRSDRNHLHDVVMQRTILELLLCFNPLWLRLGLEVVFE	660
Qy	661	KIQMQSNRDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYAETIKKHSLOKILFLLPFL	720
Db	661	KIQMQSNRDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYAETIKKHSLOKILFLLPFL	720
Qy	721	DQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGDITRELRLGYVLQHRQTF	780
Db	721	DQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGDITRELRLGYVLQHRQTF	780
Qy	781	LDEFDYAFNNLAVDLRDGVRLTRVVEVILLRDDLTQRLRVPASRLQRIFNVKLALGALG	840
Db	781	LDEFDYAFNNLAVDLRDGVRLTRVVEVILLRDDLTQRLRVPASRLQRIFNVKLALGALG	840
Qy	841	EANFQLGGDIAAQDIVDGHREKTLSELLWQLIYKFRSPKFHAAATVLQKWRRHWHLVVIQ	900
Db	841	EANFQLGGDIAAQDIVDGHREKTLSELLWQLIYKFRSPKFHAAATVLQKWRRHWHLVVIQ	900
Qy	901	RRIRHKELMRRHRAATVIQAVFRGHQMRKYVKLFKTERTQAAILQKFTRRYLAQKQLYQ	960
Db	901	RRIRHKELMRRHRAATVIQAVFRGHQMRKYVKLFKTERTQAAILQKFTRRYLAQKQLYQ	960
Qy	961	SYHSIITIQRWWRAQQQLGRQHRQRFVELREAAIFLQRIWRRRLFACKLLAAETARLQRS	1020
Db	961	SYHSIITIQRWWRAQQQLGRQHRQRFVELREAAIFLQRIWRRRLFACKLLAAETARLQRS	1020
Qy	1021	QKQQAASYYIQMQWRTYQLGRIQRHEFLRQDLIMFVQRRMRKWSMLEQRKEFQQLKRA	1080
Db	1021	QKQQAASYYIQMQWRTYQLGRIQRHEFLRQDLIMFVQRRMRKWSMLEQRKEFQQLKRA	1080
Qy	1081	AINIQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATIOMRIDRNHYHSLRKNVICLOQ	1140
Db	1081	AINIQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATIOMRIDRNHYHSLRKNVICLOQ	1140
Qy	1141	RLRAIMKMRQRENYLRLRNASILVQKRYMRQOMIQDRNAYLRTKCIINVQRRWRATL	1200
Db	1141	RLRAIMKMRQRENYLRLRNASILVQKRYMRQOMIQDRNAYLRTKCIINVQRRWRATL	1200

Qy	1201	QMRREKKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKER	1260
Db	1201	QMRREKKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKER	1260
Qy	1261	QEYLHLREVTIKLQRRFHAQKSMRFMRAKYRGTAQAVSCLQMHWRNHLLRKRENSFLQL	1320
Db	1261	QEYLHLREVTIKLQRRFHAQKSMRFMRAKYRGTAQAVSCLQMHWRNHLLRKRENSFLQL	1320
Qy	1321	RQAAITLQRRYRARLNMIKQLKSYAQLKQAAITIOTRYRAKKAMQKQVVLYQKQREAIK	1380
Db	1321	RQAAITLQRRYRARLNMIKQLKSYAQLKQAAITIOTRYRAKKAMQKQVVLYQKQREAIK	1380
Qy	1381	VQRRYRGNLEMRKQIEVYQKQRAVIRLQKWWRSIRDMRLCKAGYRRIRLSSLSIQRKWR	1440
Db	1381	VQRRYRGNLEMRKQIEVYQKQRAVIRLQKWWRSIRDMRLCKAGYRRIRLSSLSIQRKWR	1440
Qy	1441	ATVQARRQREIFLSTIRKVRMLQAFIRATLLMRQQRREFEMKRAAVVIQRRFRARCAML	1500
Db	1441	ATVQARRQREIFLSTIRKVRMLQAFIRATLLMRQQRREFEMKRAAVVIQRRFRARCAML	1500
Qy	1501	KARQDYQLIQSSVILVQRKFRANRSMKQARQEFVQLRTIAVHLQQKFRGKRLMIEQRNCF	1560
Db	1501	KARQDYQLIQSSVILVQRKFRANRSMKQARQEFVQLRTIAVHLQQKFRGKRLMIEQRNCF	1560
Qy	1561	QLLRCSMPGFQARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQ	1620
Db	1561	QLLRCSMPGFQARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQ	1620
Qy	1621	GLLDIRKRIAQLRQEAKAVNSVRCKVQEAVRFLRGRFIASDALAVLSQLDRLSRTVPHLL	1680
Db	1621	GLLDIRKRIAQLRQEAKAVNSVRCKVQEAVRFLRGRFIASDALAVLSQLDRLSRTVPHLL	1680
Qy	1681	MWCSEFMSTFCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQM	1740
Db	1681	MWCSEFMSTFCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQM	1740
Qy	1741	LLRWCDKDSEIFNTLCTLIWVFAHCPKKRKIIHDYMTNPEAIYMVRETCKKLVARKEKMKQ	1800
Db	1741	LLRWCDKDSEIFNTLCTLIWVFAHCPKKRKIIHDYMTNPEAIYMVRETCKKLVARKEKMKQ	1800
Qy	1801	NARKPPPMSTSGRYKSQKINFTPCSLPSLEPDFGIIRYSPYTFISSVYAFDTILCKLQIDM	1860
Db	1801	NARKPPPMSTSGRYKSQKINFTPCSLPSLEPDFGIIRYSPYTFISSVYAFDTILCKLQIDM	1860
Qy	1861	F 1861	
Db	1861	F 1861	

RESULT 2

ABB62757

ID ABB62757 standard; protein; 1954 AA.

XX

AC ABB62757;

XX

DT 26-MAR-2002 (first entry)

XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 15063.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL06860.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 15063; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1954 AA;

Query Match 99.7%; Score 9482; DB 4; Length 1954;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1855; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MELVWSPVLEVACKETLQLIDNRNFRKEVMIILKSKSNQPVKNPRKFPTVGKTLQLKSPT 60  
 Db 94 MELVWSPVLEVACKETLQLIDNRNFRKEVMIILKSKSNQPVKNPRKFPTVGKTLQLKSPT 153  
 Qy 61 GAGKTMKSVVSAAVQQKKRMSAAAAPPSKQTWRTAPSRPAAWAHPPPQAPLVEKNVYKT 120  
 Db 154 GAGKTMKSVVSAAVQQKKRMSAAAAPPSKQTWRTAPSRPAAWAHPPPQAPLVEKNVYKT 213  
 Qy 121 PQEEPVYISPQPRSLKENLSPMTPGNLLDVIDNLRFTPLTETRKGQATIFPDNLAAWPT 180

Db 214 PQEEPVIYISQPRSLKENLSPMTPGNLLDVIDNLRFTPLTETRGKGQATIFPDNLAAWPT 273

Qy 181 PTLKGNVKSCANDMRPRRITPDDLEDQPATNKTFDVKHSETINISLDTLDCSRIDGQPHT 240  
 |||

Db 274 PTLKGNVKSCANDMRPRRITPDDLEDQPATNKTFDVKHSETINISLDTLDCSRIDGQPHT 333

Qy 241 PLNKTTHIVHATHTRALACIHEEEGPSPPRTPTKSAIHDILKRDILKVGSPLRKYSESMKD 300  
 |||

Db 334 PLNKTTHIVHATHTRALACIHEEEGPSPPRTPTKSAIHDILKRDILKVGSPLRKYSESMKD 393

Qy 301 LSLSPQTKYAIQGSMPNLNEMKIRSIEQNRYQEQQIQIKAKDLNSSSSSEASLAGQQE 360  
 |||

Db 394 LSLSPQTKYAIQGSMPNLNEMKIRSIEQNRYQEQQIQIKAKDLNSSSSSEASLAGQQE 453

Qy 361 FLFNHSEILAQSSRNLHEVGRKSVKGSVPKNPHKRRSHELFSFDAPSNESELYRNETVAI 420  
 |||

Db 454 FLFNHSEILAQSSRNLHEVGRKSVKGSVPKNPHKRRSHELFSFDAPSNESELYRNETVAI 513

Qy 421 SPPKKQRVEDTTLPRSAAPANASARSSSAHAWPHAQSKKFKLAQTMSLMKKPATPRKVRD 480  
 |||

Db 514 SPPKKQRVEDTTLPRSAAPANASARSSSAHAWPHAQSKKFKLAQTMSLMKKPATPRKVRD 573

Qy 481 TSIQPSVKLYDSELYMQTCINPDFAATTTIDPFLASTMYLDEQAVDRHQADFKKWLNAL 540  
 |||

Db 574 TSIQPSVKLYDSELYMQTCINPDFAATTTIDPFLASTMYLDEQAVDRHQADFKKWLNAL 633

Qy 541 VSIPADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNYLTKYRLETLRKAAVELFF 600  
 |||

Db 634 VSIPADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNYLTKYRLETLRKAAVELFF 693

Qy 601 SEQMRLPCSKVAVYVVKQALRIRSDRNHLHDVVMQRTILELLLCFNPLWLRLGLEVVFE 660  
 |||

Db 694 SEQMRLPCSKVAVYVVKQALRIRSDRNHLHDVVMQRTILELLLCFNPLWLRLGLEVVFE 753

Qy 661 KIQQSNRDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYAETIKKHSQKILFLLPFL 720  
 |||

Db 754 KIQQSNRDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYAETIKKHSQKILFLLPFL 813

Qy 721 DQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGDITRELRLGYVLQHRQTF 780  
 |||

Db 814 DQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGDITRELRLGYVLQHRQTF 873

Qy 781 LDEFDYAFNNLAVDLRDGVRRLTRVVEVILLRDDLTQRLRVPASRLQRI FNVKLALGALG 840  
 |||

Db 874 LDEFDYAFNNLAVDLRDGVRRLTRVMEVILLRDDLTQRLRVPASRLQRI FNVKLALGALG 933

Qy 841 EANFQLGGDIAAQDIVDGHREKTL SLLWQLIYKFRSPKFHAAATVLQKWRRHWLHVVIQ 900  
 |||

Db 934 EANFQLGGDIAAQDIVDGHREKTL SLLWQLIYKFRSPKFHAAATVLQKWRRHWLHVVIQ 993

Qy 901 RRIRHKELMRRHRAATVIQAVFRGHQMRKYVKLFKKTERTQAAIILQKFTRRYLAQKQLYQ 960  
 |||

Db 994 RRIRHKELMRRHRAATVIQAVFRGHQMRKYVKLFKKTERTQAAIILQKFTRRYLAQKQLYQ 1053

Qy 961 SYHSIITIQRWWRAQQQLGRQHRQRFVELREAAIFLQRIWRRRLFAKKLLAAAEARLQRS 1020  
 |||

Db 1054 SYHSIITIQRWWRAQQQLGRQHRQRFVELREAAIFLQRIWRRRLFAKKLLAAAEARLQRS 1113



Qy	1021	QKQQAASYYIQMWRTYQLGRIQRHEFLRQRD LIMFVQRRMRSKWSMLEQRKEFQQLKRA	1080
		:     :	
Db	1114	QKQQAASYYIQMQRSYQLGRIQRQQFLRQRD LIMFVQRRMRSKWSMLEQRKEFQQLKRA	1173
Qy	1081	AINIQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATI QMRIDRNHYYS LRKNVICLQQ	1140
Db	1174	AINIQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATI QMRIDRNHYYS LRKNVICLQQ	1233
Qy	1141	RLRAIMKMREQRENYLRLRNAS ILVQKRYMRQOMIQDRNAYLRT RKCI INVQRRWRATL	1200
Db	1234	RLRAIMKMREQRENYLRLRNAS ILVQKRYMRQOMIQDRNAYLRT RKCI INVQRRWRATL	1293
Qy	1201	QMRREKKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKER	1260
Db	1294	QMRREKKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKER	1353
Qy	1261	QEYLHLREVTIKLQRRFHAQKSMRFMRACYRGTOAAVSC LQMHWRNHLLRKRENSFLQL	1320
Db	1354	QEYLHLREVTIKLQRRFHAQKSMRFMRACYRGTOAAVSC LQMHWRNHLLRKRENSFLQL	1413
Qy	1321	RQAAITLQRRYRARNMIKQLKSYAQLKQAAITIQTRYRAKKAMQKQVVLYQKQREAIK	1380
Db	1414	RQAAITLQRRYRARNMIKQLKSYAQLKQAAITIQTRYRAKKAMQKQVVLYQKQREAIK	1473
Qy	1381	VQRRYRGNLEMRKQIEVYQKQRAVIRLQKWWRSIRDMLCKAGYRRIRLSSLSIQRKWR	1440
Db	1474	VQRRYRGNLEMRKQIEVYQKQRAVIRLQKWWRSIRDMLCKAGYRRIRLSSLSIQRKWR	1533
Qy	1441	ATVQARRQREIFLSTIRKVRMLQAFIRATLLMRQORREFEMKRRAAVVIQRRFRARCAML	1500
Db	1534	ATVQARRQREIFLSTIRKVRMLQAFIRATLLMRQORREFEMKRRAAVVIQRRFRARCAML	1593
Qy	1501	KARQDYQLIQSSVILVQRKFRANRSMKQARQEFVQLRTIAVHLQQKFRGKRLMIEQRNCF	1560
Db	1594	KARQDYQLIQSSVILVQRKFRANRSMKQARQEFVQLRTIAVHLQQKFRGKRLMIEQRNCF	1653
Qy	1561	QLLRCSMPGFQARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQ	1620
Db	1654	QLLRCSMPGFQARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQ	1713
Qy	1621	GLLDIRKRIAQLRQEAKAVNSVRCKVQEA VRFLRGRFIASDALAVLSQLDRLSRTVP HLL	1680
		:	
Db	1714	GLLDIRKRIAQLRQEAKAVNSVRCKVQEA VRFLRGRFIASDALAVLSRLDRLSRTVP HLL	1773
Qy	1681	MWCSEFMSTFCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQM	1740
Db	1774	MWCSEFMSTFCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQM	1833
Qy	1741	LLRWCDKDSEIFNTLCTLIWVFAHCPKKRKIIHDYMTNPEAIYMVRETKKLVARKEKMKQ	1800
Db	1834	LLRWCDKDSEIFNTLCTLIWVFAHCPKKRKIIHDYMTNPEAIYMVRETKKLVARKEKMKQ	1893
Qy	1801	NARKPPPMSTSGRYKSQKINFTPCSLPSLEPDFGIIRYSPYTFISSVYAFDTILCKLQIDM	1860
Db	1894	NARKPPPMSTSGRYKSQKINFTPCSLPSLEPDFGIIRYSPYTFISSVYAFDTILCKLQIDM	1953

Qy 1861 F 1861  
|  
Db 1954 F 1954

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 22, 2005, 23:31:55 ; Search time 61 Seconds  
(without alignments)  
2935.400 Million cell updates/sec

Title: US-09-914-698-1  
Perfect score: 9514  
Sequence: 1 MELVWSPVLEVACKETLQLI.....FISSVYAFDTILCKLQIDMF 1861

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9514	100.0	1861	2	T13845	microtubule-associ
2	361.5	3.8	1186	2	T19957	hypothetical prote
3	324.5	3.4	1088	2	T05464	hypothetical prote
4	324	3.4	1407	1	S28589	trichohyalin - rab
5	323.5	3.4	1549	1	A40691	trichohyalin - she
6	320	3.4	1830	1	S19188	myosin-V - chicken
7	314	3.3	1909	2	A45592	liver stage antige
8	304	3.2	1828	2	B59254	mysoin heavy chain
9	301	3.2	4574	2	G02520	plectin - human
10	301	3.2	4687	1	A39638	plectin - rat

11	300.5	3.2	1375	2	T05200	myosin heavy chain
12	300.5	3.2	1855	2	A59254	myosin heavy chain
13	300	3.2	1374	2	D85390	myosin-like protei
14	294	3.1	4684	2	A59404	plectin [imported]
15	284.5	3.0	1805	1	A64224	hypothetical prote
16	277	2.9	1818	1	S73852	hypothetical prote
17	277	2.9	1898	1	A45973	trichohyalin - hum
18	271.5	2.9	1853	1	A46761	myosin heavy chain
19	267.5	2.8	2663	1	S28261	centromere protein
20	264	2.8	2442	2	T08621	centrosome associa
21	260	2.7	2245	2	T18278	myosin heavy chain
22	254.5	2.7	1837	2	T41023	probable nuclear p
23	252	2.6	1313	2	A48467	myosin heavy chain
24	251.5	2.6	1792	2	T13939	myosin V - fruit f
25	250	2.6	2101	2	A42184	nuclear mitotic ap
26	246.5	2.6	2022	2	T43214	ovtl protein - nem
27	244	2.6	1846	2	A59289	myr 6, unconventio
28	243.5	2.6	812	2	A53016	myosin heavy chain
29	243.5	2.6	1410	1	A57013	early endosome ant
30	239.5	2.5	1940	2	A59287	myosin heavy chain
31	235	2.5	2253	2	T30336	nuclear/mitotic ap
32	235	2.5	3187	2	JC5837	364K Golgi complex
33	233.5	2.5	1643	2	T07961	myosin heavy chain
34	231	2.4	968	2	T00322	hypothetical prote
35	230.5	2.4	1211	2	C83110	probable exonuclea
36	228.5	2.4	1679	2	S48385	hypothetical prote
37	228	2.4	1957	2	A59294	skeletal myosin -
38	227.5	2.4	3259	1	A56539	giantin - human
39	227	2.4	1957	2	A45627	myosin heavy chain
40	227	2.4	3678	2	S28916	dystrophin - mouse
41	226	2.4	2346	2	T13829	Tpr homolog - frui
42	226	2.4	2626	2	T31099	myosin-RhoGAP prot
43	225.5	2.4	3225	2	I52300	giantin - human
44	225	2.4	3685	1	A27605	dystrophin, muscle
45	224.5	2.4	1927	2	A59236	embryonic muscle m

#### ALIGNMENTS

RESULT 1  
 T13845  
 microtubule-associated protein - fruit fly (*Drosophila melanogaster*)  
 C;Species: *Drosophila melanogaster*  
 C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C;Accession: T13845  
 R;Saunders, R.D.; Avides, M.C.; Howard, T.; Gonzalez, C.; Glover, D.M.  
 J. Cell Biol. 137, 881-890, 1997  
 A;Title: The *Drosophila* gene abnormal spindle encodes a microtubule-associated protein that associates with the polar regions of the mitotic spindle.  
 A;Reference number: Z17792; MUID:97296495; PMID:9151690  
 A;Accession: T13845  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-1861 <SAU>  
 A;Cross-references: UNIPROT:O01401; EMBL:U95171; NID:g1930121; PID:g1930122;  
 PIDN:AAB51540.1

C;Genetics:  
A;Gene: asp  
A;Cross-references: FlyBase:FBgn0000140  
C;Function:  
A;Description: is required for the normal function of the mitotic spindle

Query Match 100.0%; Score 9514; DB 2; Length 1861;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MELVWSPVLEVACKETLQLIDNRNFRKEVMIILKSKSNQPVKNPRKFPTVGKTLQLKSPT 60
      |||
Db      1 MELVWSPVLEVACKETLQLIDNRNFRKEVMIILKSKSNQPVKNPRKFPTVGKTLQLKSPT 60

Qy     61 GAGKTMKSVVSAAVQQKKRMSAAAAPPSKQTWRVTAPSRAAWAHPPPQAPLVEKNVYKT 120
      |||
Db     61 GAGKTMKSVVSAAVQQKKRMSAAAAPPSKQTWRVTAPSRAAWAHPPPQAPLVEKNVYKT 120

Qy    121 PQEEPVIYISQPRSLKENLSPMTPGNLLDVIDNLRFTPLTETRKGQATIFPDNLAAWPT 180
      |||
Db    121 PQEEPVIYISQPRSLKENLSPMTPGNLLDVIDNLRFTPLTETRKGQATIFPDNLAAWPT 180

Qy    181 PTLKGNVKSCANDMRPRRITPDDLEDQPATNKTFDVKHSETINISLDTLDCSRIDGQPHT 240
      |||
Db    181 PTLKGNVKSCANDMRPRRITPDDLEDQPATNKTFDVKHSETINISLDTLDCSRIDGQPHT 240

Qy    241 PLNKTITIVHATHTRALACIHEEEGPSPPRTPTKSAIHDLKRDIKLVGSPLRKYSESMKD 300
      |||
Db    241 PLNKTITIVHATHTRALACIHEEEGPSPPRTPTKSAIHDLKRDIKLVGSPLRKYSESMKD 300

Qy    301 LSLSPQTKYAIQGSMPNLNEMKIRSIEQNRYQQEQIQIKAKDLNSSSSSEASLAGQQE 360
      |||
Db    301 LSLSPQTKYAIQGSMPNLNEMKIRSIEQNRYQQEQIQIKAKDLNSSSSSEASLAGQQE 360

Qy    361 FLFNHSEILAQSSRFNLHEVGRKSVKGSVPKNPHKRRSHELFSFDAPSNESELYRNETVAI 420
      |||
Db    361 FLFNHSEILAQSSRFNLHEVGRKSVKGSVPKNPHKRRSHELFSFDAPSNESELYRNETVAI 420

Qy    421 SPPKKQRVEDTTLPRSAAPANASARSSSAHAWPHAQSKKFKLAQTMSLMKKPATPRKVRD 480
      |||
Db    421 SPPKKQRVEDTTLPRSAAPANASARSSSAHAWPHAQSKKFKLAQTMSLMKKPATPRKVRD 480

Qy    481 TSIQPSVKLYDSELYMQTCINPDFAATTTIDPFLASTMYLDEQAVDRHQADFKKWLNAL 540
      |||
Db    481 TSIQPSVKLYDSELYMQTCINPDFAATTTIDPFLASTMYLDEQAVDRHQADFKKWLNAL 540

Qy    541 VSIPADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNYLTKYRLETLRKAAVELFF 600
      |||
Db    541 VSIPADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNYLTKYRLETLRKAAVELFF 600

Qy    601 SEQMRLPCSKVAVYVVKQALRIRSDRNHLHDVVMQRTILELLLCFNPLWLRGLGVVFFE 660
      |||
Db    601 SEQMRLPCSKVAVYVVKQALRIRSDRNHLHDVVMQRTILELLLCFNPLWLRGLGVVFFE 660

Qy    661 KIQMQSNRDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYAETIKKHSLOKILFLLPFL 720
      |||
Db    661 KIQMQSNRDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYAETIKKHSLOKILFLLPFL 720
```

Qy	721	DQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGDITRELRRLG YVLQHRQTF	780
Db	721	DQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGDITRELRRLG YVLQHRQTF	780
Qy	781	LDEFDYAFNNLAVDLRDGVRLTRVVEVILLRDDLTRQLRVPAISRLQRI FNVKLALGALG	840
Db	781	LDEFDYAFNNLAVDLRDGVRLTRVVEVILLRDDLTRQLRVPAISRLQRI FNVKLALGALG	840
Qy	841	EANFQLGGDIAAQDIVDGHREKTL SLLWQLIYKFRSPKFHAAATVLQKWWRRHLHVVIQ	900
Db	841	EANFQLGGDIAAQDIVDGHREKTL SLLWQLIYKFRSPKFHAAATVLQKWWRRHLHVVIQ	900
Qy	901	RRIRHKELMRRHRAATVIAVFRGHQMRKYVKLFKTERTQAAIILQKFTRRYLAQKQLYQ	960
Db	901	RRIRHKELMRRHRAATVIAVFRGHQMRKYVKLFKTERTQAAIILQKFTRRYLAQKQLYQ	960
Qy	961	SYHSIITIQRWWRAQQQLGRQHRQRFVELREAAIFLQRIWRRRLFAKKLLAAETARLQRS	1020
Db	961	SYHSIITIQRWWRAQQQLGRQHRQRFVELREAAIFLQRIWRRRLFAKKLLAAETARLQRS	1020
Qy	1021	QKQQAASYSIQMWRTYQLGRIQRHEFLRQDLIMFVQRRMRSKWSMLEQRKEFQQLKRA	1080
Db	1021	QKQQAASYSIQMWRTYQLGRIQRHEFLRQDLIMFVQRRMRSKWSMLEQRKEFQQLKRA	1080
Qy	1081	AINIQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATI QMRIDRNHYYSRKNVICLQQ	1140
Db	1081	AINIQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATI QMRIDRNHYYSRKNVICLQQ	1140
Qy	1141	RLRAIMKMREQRENYLRLRNASILVQKRYMRQOMIQDRNAYLRTKCIINVQRRWRATL	1200
Db	1141	RLRAIMKMREQRENYLRLRNASILVQKRYMRQOMIQDRNAYLRTKCIINVQRRWRATL	1200
Qy	1201	QMRERKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKER	1260
Db	1201	QMRERKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKER	1260
Qy	1261	QEYLHLREVTIKLQRRFHAQKSMRFMRKYRGTQAAVSCLOMHWNRHLLRKRENSFLQL	1320
Db	1261	QEYLHLREVTIKLQRRFHAQKSMRFMRKYRGTQAAVSCLOMHWNRHLLRKRENSFLQL	1320
Qy	1321	RQAAITLQRRYRARLNMIKQLKSYAQLKQAAITI QTRYRAKKAMQKQVVLYQKQREAIK	1380
Db	1321	RQAAITLQRRYRARLNMIKQLKSYAQLKQAAITI QTRYRAKKAMQKQVVLYQKQREAIK	1380
Qy	1381	VQRRYRGNLEMRKQIEVYQKQRAVIRLQKWWRSIRDMLCKAGYRRIRLSSLSIQRKWR	1440
Db	1381	VQRRYRGNLEMRKQIEVYQKQRAVIRLQKWWRSIRDMLCKAGYRRIRLSSLSIQRKWR	1440
Qy	1441	ATVQARRQREIFLSTIRKVRMLQAFIRATLLMRQQRREFEMKRRAAVVIQRRFRARCAML	1500
Db	1441	ATVQARRQREIFLSTIRKVRMLQAFIRATLLMRQQRREFEMKRRAAVVIQRRFRARCAML	1500
Qy	1501	KARQDYQLIQSSVILVQRKFRANRSMKQARQEFVQLRTIAVHLQQKFRGKRLMIEQRNCF	1560
Db	1501	KARQDYQLIQSSVILVQRKFRANRSMKQARQEFVQLRTIAVHLQQKFRGKRLMIEQRNCF	1560

Qy 1561 QLLRCSMPGFQARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQ 1620  
 |||  
 Db 1561 QLLRCSMPGFQARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQ 1620

Qy 1621 GLLDIRKRIAQLRQEAKAVNSVRCKVQEAVRFLRGRFIASDALAVLSQLDRLSRTVPHLL 1680  
 |||  
 Db 1621 GLLDIRKRIAQLRQEAKAVNSVRCKVQEAVRFLRGRFIASDALAVLSQLDRLSRTVPHLL 1680

Qy 1681 MWCSEFMSTFCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQM 1740  
 |||  
 Db 1681 MWCSEFMSTFCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQM 1740

Qy 1741 LLRWCDKDSEIFNTLCTLIWVFAHCPKKRKIIHDYMTNPEAIYMVRETKKLVARKEKMKQ 1800  
 |||  
 Db 1741 LLRWCDKDSEIFNTLCTLIWVFAHCPKKRKIIHDYMTNPEAIYMVRETKKLVARKEKMKQ 1800

Qy 1801 NARKPPPMSTSGRYKSQKINFTPCSLPSLEPDFGIIRYSPYTFISSVYAFDTILCKLQIDM 1860  
 |||  
 Db 1801 NARKPPPMSTSGRYKSQKINFTPCSLPSLEPDFGIIRYSPYTFISSVYAFDTILCKLQIDM 1860

Qy 1861 F 1861  
 |  
 Db 1861 F 1861

## RESULT 2

T19957

hypothetical protein C45G3.1 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T19957

R;Barlow, K.

submitted to the EMBL Data Library, March 1997

A;Reference number: Z19203

A;Accession: T19957

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1186 <WIL>

A;Cross-references: UNIPROT:O17666; EMBL:Z92780; PIDN:CAB07174.1; GSPDB:GN00019;  
 CESP:C45G3.1

A;Experimental source: clone C45G3

C;Genetics:

A;Gene: CESP:C45G3.1

A;Map position: 1

A;Introns: 21/1; 255/2; 363/2; 575/3; 893/3; 1017/2; 1042/1

C;Superfamily: *Caenorhabditis elegans* hypothetical protein C45G3.1

Query Match 3.8%; Score 361.5; DB 2; Length 1186;

Best Local Similarity 18.6%; Pred. No. 6.5e-12;

Matches 244; Conservative 216; Mismatches 480; Indels 375; Gaps 47;

Qy 458 KKFkLAQTMSLMKK---PATPRKVRDTSIQPSVKLYDSELYMQTCINPDFFAATTTIDPF 514  
 :| | | | || || | :| : : | :| :| : : |  
 Db 21 EKRLLDQVKSNTKKIDLRATERAFLESS-----PTSMNMRTPLNPS-ISSSVSDSPI 71

Qy 515 LASTMYLDEQAVDRHQADFKKWLNALVSIPADLDADLNNKIDVGKLFNEV----- 564  
 |: ||:| :: | | : :: |:| ::|:| |

Db 72 LS----YDEKA-NKQIIALATWCNTM-----MELDVSEEMDLGESKAEACRNIQKMLKK 120

Qy 565 RNKELVVAPTKEEQSMNY---LTKYRLETLRKAAVELFFSEQMRLPCSKVAVYVVKQALR 621  
|: | |:| | | :| :| | : : :| :

Db 121 RSDTSEVENTQENARRRYQRIFEKNDPEVVKKCKQLLDDSGMD---ASIKDLLSKNNVA 177

Qy 622 IRSDRNHLHDVVMQRTILELLLCFNPLWLRLGLEVVFGKIQMQSNRDIVGLSTFILNRL 681  
|| : :| :| |:| | |:| ||: || :| :| | : || | |: :

Db 178 IRKEHAVYNDIGLQTTLLHTFLSFHPAWLKTALAEAFNTRIDAQPKHLMKKLSQFFLDLV 237

Qy 682 FRN--KCEEQRYSKAY---TLTEEYAETIKKHSLOKILFLLPFLDQAKQKRIVKHNPCLF 736  
| | : : : : :| | | : | | : :| | : : :| :

Db 238 FSNPTMLKNKKFAQGSGKPIITEAGKEALHKHFLSVSMKLMFLIETAHTRVIPNLTRIF 297

Qy 737 VKKSPHKETKDILLRFSSELL-ANIGDITRELRLGYVLQHRQTFLEFDYAFNNLAVDL 795  
| | | : : || : : : : :| :| :| :| :| :| :

Db 298 TKSSHFNCLDDVSELTKELLTGSSATFKKAFKVGFIPTYRQSFENYDYQAKGFS-DF 356

Qy 796 RDGVRLTRVVEVI--LLRDDLTRQLRVPAPISRLQRI FNVKLALGALGEANFQLG---GDI 850  
||: | : :| : : : || || | : :| ||: | : || : :

Db 357 SDGLILAKLLETVGEMPHGQILLKL RDPAGDRIRKIGNVKIVLQEMS----SLGVPTDNV 412

Qy 851 AAQDIVDGHREKTL SLLWQLIYKFRSPKFHAAATVLQKWWRRHWLHVVIQRRIR-----H 905  
|: || | : : ||: || :| : | :| :

Db 413 NAESIVGGKKDEILSILWAI-----GVRVAKEQRIKVTRVSE 450

Qy 906 KELMRRHRAA-----TVIQAVFRGHQMRKYVKLFKTERTQAAILQKFTRRYLAQKQ 957  
: : : |:| : : | || : : : :| | :

Db 451 ERTPKKRRSAVHDDMSSEVLKMKKIYGRQME--IEVMDLDSLSDGCLLAKLWTTFGTNST 508

Qy 958 LYQSYHSIITIQRWWRAQQLGRQHRQRFVELREAAIFLQR--IWRRLFAKKLLAAAETA 1015  
| | : | : : :| : | : :| || | |

Db 509 PIQDYDG---LSLW-----EKVVSVAELELCIQGLDQNMALFVKMFL-----E 549

Qy 1016 RLQRSQKQQAASYIQMWRTYQLGRIQRHE-----FL-----RQRDLIMF--- 1056  
|| | | : || | : | :|| | : | | : |

Db 550 RLGMIOQLNEKATKIQRMWKAY----VQRKNTPKLYFIVQQLLADSSIPRNSVSPFSNN 605

Qy 1057 ----VQRRMRKWSMLEQRKEFQQLKRAAINIQQRWRAKLSMRKCNADYLALRSSVLKVQ 1112  
| | | : :| :| | : : | : :| :| :| :

Db 606 VTFTVPRTPRN--NILTERPSLSQIPSS-----RQSMDSFTNDATFTVSRDSIESMN 655

Qy 1113 AYRKA-----TIQMRIDRNHYYSRKNVICLQQLRAIMKMREQRENYLRLRNASI 1163  
:| || | | : :| : :| :| :| :

Db 656 KMQKTPLRGTFTRKTIAMVIEEEDDSENNETVVPSTLKKRTVVRMEHNAEVF----- 707

Qy 1164 LVQKRYRMQOMIQDRNAYLRTRKCIINVQRRWRATLQMRERKNYLHLQTTTKRIQIKF 1223  
|:| | | : : | : | :| : | : :

Db 708 -----REQDEDEN-----QDKDTVAPSAE-----NLDSPPSDIPLT 740

Qy 1224 RAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKERQEYLHLREVTIKLQRRFHAQKSM 1283  
: | | || : | |

Db 741 LSSIPASQSAIFLQDSE-----TGKEM 763

Qy 1284 RFMRAKYRGT---QAAVSCLQMHWRNHLLRKRERNSFLQLRQAAITLQRRYRARNMIKQ 1340  
:|: | :| :| : : | :| :| :| :

Db 764 HVPKAEDVGVVVLEASDSPVALEGNN-----EASYDGQKIENLETFEIKE 808

```

Qy      1341 LKSYAQLKQAAITIQTRYRAKKAMQKQVVLQKQREAIKQVRRYRGNLEMRKQIEVYQK 1400
      | : | : | : : : | : : : | : :
Db      809 GKTQEDLPSPMDPTQTSGSPLEFRMTTEEQRLEMLFQSLSEDQKNFVKTNLSVSIE 868

Qy      1401 QRQAVIRLQKWRSIRDMLRCKAGYRRIRLSSLSIQKWRATVQARRQREIFLSTIRKVR 1460
      | : | | : : : | : : |
Db      869 DDANTPELRRILRQTRELK-----RKQQEI-----AR 895

Qy      1461 LMQAFIRATLLMRQQRREFEMKRRAAVVIQRRFRARCAMLKARQDYQLIQSSVILVQRKF 1520
      : | | : | : | : | : | : | : | :
Db      896 KLGNIERNALAVRDGGEDSSDSRSDA-----GHDVAILHGDDSQLFENSMQLDQK-- 945

Qy      1521 RANRSMKQARQEFVQLRTIAVHLQQKFRGKRLMIEQRNCFQLLRCSMPGFQARARGFMAR 1580
Db      946 ----- 945

Qy      1581 KRFOALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQGLLDIRKRIAQ----LRQEA 1636
      | | | : | : | : | : | : | : | : | : | : | :
Db      946 ---SQLQNDTQILENKKAQVVIQKMIRGFIARRKFQME-ISNIRNRMIQYNHILAQED 1001

Qy      1637 KAV-----NSVRCKVQEA VRFLRGRFIASDALAVL-----SQLDRLSRTVPHLL 1680
      : : | | : : | : | : : | : : | : | :
Db      1002 EQIGIEEMEDKSVEAKLKKCA--LHG--LTNDNLHVHVHAATVIDRVTDLVPSLL 1052

```

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 22, 2005, 23:22:29 ; Search time 244 Seconds  
(without alignments)  
3905.653 Million cell updates/sec

Title: US-09-914-698-1  
Perfect score: 9514  
Sequence: 1 MELVWSPVLEVACKETLQLI.....FISSVYAFDTILCKLQIDMF 1861

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a



score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	%		Query Match Length	DB	ID	Description
	Score	Match				
1	9482	99.7	1954	1	ASP_DROME	Q9vc45 drosophila
2	3095	32.5	1538	2	Q7QAG9	Q7qag9 anopheles g
3	1407.5	14.8	3452	1	ASPM_CANFA	P62286 canis famil
4	1392	14.6	3371	1	ASPM_BOVIN	P62285 bos taurus
5	1385	14.6	3469	1	ASPM_SAIBB	P62296 saimiri bol
6	1374	14.4	3122	1	ASPM_MOUSE	Q8cj27 mus musculu
7	1373	14.4	3477	1	ASPM_PANTR	P62293 pan troglod
8	1366.5	14.4	3476	1	ASPM_GORG0	P62289 gorilla gor
9	1363	14.3	3471	1	ASPM_PONPY	P62294 pongo pygma
10	1362.5	14.3	3473	1	ASPM_AOTVO	P62283 aotus vocif
11	1360	14.3	3477	1	ASPM_HUMAN	Q8izt6 homo sapien
12	1360	14.3	3477	1	ASPM_HYLLA	P62290 hylobates l
13	1359	14.3	3374	1	ASPM_SHEEP	P62297 ovis aries
14	1356.5	14.3	3461	1	ASPM_FELCA	P62288 felis silve
15	1353.5	14.2	3475	2	Q69AX5	Q69ax5 cercopithec
16	1348.5	14.2	3479	1	ASPM_MACMU	P62292 macaca mula
17	1343	14.1	3476	1	ASPM_MACFA	P62291 macaca fasc
18	1330.5	14.0	3477	1	ASPM_COLGU	P62287 colobus gue
19	744.5	7.8	1564	1	ASPM_ATEGE	P62284 ateles geof
20	729.5	7.7	1527	1	ASPM_SAGLB	P62295 saguinus la
21	370.5	3.9	1110	2	Q84Z28	Q84z28 oryza sativ
22	361.5	3.8	1186	2	O17666	O17666 caenorhabdi
23	324.5	3.4	1088	2	O49705	O49705 arabidopsis
24	324	3.4	1407	1	TRHY_RABIT	P37709 oryctolagus
25	323.5	3.4	1549	1	TRHY_SHEEP	P22793 ovis aries
26	317.5	3.3	1829	1	MY5A_CHICK	Q02440 gallus gall
27	314	3.3	1909	2	Q25893	Q25893 plasmodium
28	303.5	3.2	2779	1	LVA_DROME	Q8mssl drosophila
29	301.5	3.2	1855	1	MY5A_HUMAN	Q9y4i1 homo sapien
30	301	3.2	4515	2	Q6S379	Q6s379 homo sapien
31	301	3.2	4525	2	Q6S381	Q6s381 homo sapien
32	301	3.2	4533	2	Q6S382	Q6s382 homo sapien
33	301	3.2	4547	2	Q6S376	Q6s376 homo sapien
34	301	3.2	4547	2	Q6S378	Q6s378 homo sapien
35	301	3.2	4551	2	Q6S377	Q6s377 homo sapien
36	301	3.2	4574	2	Q6S383	Q6s383 homo sapien
37	301	3.2	4684	2	Q6S380	Q6s380 homo sapien
38	301	3.2	4687	1	PLE1_RAT	P30427 rattus norv
39	300	3.2	1374	2	Q9SMY9	Q9smy9 arabidopsis
40	295.5	3.1	1828	1	MY5A_RAT	Q9qyf3 rattus norv
41	294	3.1	4684	1	PLE1_HUMAN	Q15149 homo sapien
42	292	3.1	4451	2	Q6S3A1	Q6s3a1 rattus norv
43	292	3.1	4513	2	Q6S399	Q6s399 rattus norv
44	292	3.1	4523	2	Q6S3A3	Q6s3a3 rattus norv
45	292	3.1	4536	2	Q6S3A4	Q6s3a4 rattus norv

# ALIGNMENTS

# RESULT 1

## ASP\_DROME

ID ASP\_DROME STANDARD; PRT; 1954 AA.  
AC Q9VC45; O01401; Q8SX66;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE Abnormal spindle protein.  
GN Name=asp; ORFNames=CG6875;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).

RN [2]  
 RP GENOME REANNOTATION.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review."  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [3]  
 RP SEQUENCE OF 94-1954 FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND  
 RP DEVELOPMENTAL STAGE.  
 RC STRAIN=Oregon-R;  
 RX PubMed=9151690; DOI=10.1083/jcb.137.4.881;  
 RA Saunders R.D.C., do Carmo Avides M., Howard T.I.A., Gonzalez C.,  
 RA Glover D.M.;  
 RT "The *Drosophila* gene abnormal spindle encodes a novel microtubule-  
 RT associated protein that associates with the polar regions of the  
 RT mitotic spindle."  
 RL J. Cell Biol. 137:881-890(1997).  
 RN [4]  
 RP SEQUENCE OF 700-1954 FROM N.A.  
 RC STRAIN=Berkeley; TISSUE=Embryo;  
 RX MEDLINE=22426066; PubMed=12537569;  
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
 RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,  
 RA Rubin G.M., Celniker S.E.;  
 RT "A *Drosophila* full-length cDNA resource."  
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
 RN [5]  
 RP FUNCTION, AND SUBCELLULAR LOCATION.  
 RX PubMed=10073938; DOI=10.1126/science.283.5408.1733;  
 RA do Carmo Avides M., Glover D.M.;  
 RT "Abnormal spindle protein, Asp, and the integrity of mitotic  
 RT centrosomal microtubule organizing centers."  
 RL Science 283:1733-1735(1999).  
 RN [6]  
 RP FUNCTION.  
 RX PubMed=15242765; DOI=10.1016/j.yexcr.2004.03.054;  
 RA Riparbelli M.G., Massarelli C., Robbins L.G., Callaini G.;  
 RT "The abnormal spindle protein is required for germ cell mitosis and  
 RT oocyte differentiation during *Drosophila* oogenesis."  
 RL Exp. Cell Res. 298:96-106(2004).  
 CC -!- FUNCTION: Required to maintain the structure of the centrosomal  
 CC microtubule-organizing center (MTOC) during mitosis. May have a  
 CC preferential role in regulating neurogenesis. Required for germ  
 CC cell mitosis and oocyte differentiation.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear. During interphase  
 CC in syncytial embryos distribution is cytoplasmic. On entering  
 CC mitosis, moves to polar regions of the spindle immediately  
 CC surrounding the centrosome. At telophase, migrates to microtubules  
 CC on the spindle side of both daughter nuclei. The nuclear-  
 CC cytoplasmic distribution could be regulated by the availability of

CC calmodulin.  
 CC -!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically in  
 CC embryos.  
 CC -!- SIMILARITY: Contains 1 calponin-homology (CH) domain.  
 CC -!- SIMILARITY: Contains 5 IQ domains.  
 CC -!- CAUTION: Ref.4 sequence differs from that shown due to intron  
 CC retention.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE003749; AAF56330.3; -.  
 DR EMBL; U95171; AAB51540.1; -.  
 DR EMBL; AY094825; AAM11178.1; ALT\_SEQ.  
 DR PIR; T13845; T13845.  
 DR FlyBase; FBgn0000140; asp.  
 DR GO; GO:0005875; C:microtubule associated complex; IDA.  
 DR GO; GO:0008017; F:microtubule binding; IDA.  
 DR GO; GO:0004672; F:protein kinase activity; IDA.  
 DR InterPro; IPR001715; Calponin-like.  
 DR InterPro; IPR000048; IQ\_region.  
 DR Pfam; PF00307; CH; 1.  
 DR Pfam; PF00612; IQ; 21.  
 DR SMART; SM00033; CH; 1.  
 DR SMART; SM00015; IQ; 16.  
 DR PROSITE; PS50021; CH; 1.  
 DR PROSITE; PS50096; IQ; 5.  
 KW Calmodulin-binding; Cell division; Coiled coil; Microtubule; Mitosis;  
 KW Nuclear protein; Oogenesis; Repeat.  
 FT DOMAIN 836 968 CH.  
 FT DOMAIN 1004 1033 IQ 1.  
 FT DOMAIN 1386 1415 IQ 2.  
 FT DOMAIN 1467 1496 IQ 3.  
 FT DOMAIN 1614 1641 Coiled coil (Potential).  
 FT DOMAIN 1656 1687 IQ 4.  
 FT DOMAIN 1690 1721 IQ 5.  
 FT DOMAIN 1063 1749 Arg-rich.  
 FT CONFLICT 811 811 L -> P (in Ref. 3).  
 FT CONFLICT 898 898 M -> V (in Ref. 3).  
 FT CONFLICT 1129 1129 S -> T (in Ref. 3).  
 FT CONFLICT 1138 1139 QQ -> HE (in Ref. 3).  
 FT CONFLICT 1761 1761 R -> Q (in Ref. 1).  
 SQ SEQUENCE 1954 AA; 230178 MW; 4912B4E20CA9E659 CRC64;

Query Match 99.7%; Score 9482; DB 1; Length 1954;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1855; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MELVWSPVLEVACKETLQLIDNRNFRKEVMIILKSKSNQPVKNPRKFPTVGKTLQLKSPT 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 94 MELVWSPVLEVACKETLQLIDNRNFRKEVMIILKSKSNQPVKNPRKFPTVGKTLQLKSPT 153

Qy 61 GAGKTMKSVVSAAVQQKKRMSAAAAPPSKQTWRVTAPSRPAAWAHPPPQAPLVEKNVYKT 120  
 |||  
 Db 154 GAGKTMKSVVSAAVQQKKRMSAAAAPPSKQTWRVTAPSRPAAWAHPPPQAPLVEKNVYKT 213

Qy 121 PQEEPVIISPQPRSLKENLSPMTPGNLLDVIDNLRFTPLTETRGKGQATIFPDNLAAWPT 180  
 |||  
 Db 214 PQEEPVIISPQPRSLKENLSPMTPGNLLDVIDNLRFTPLTETRGKGQATIFPDNLAAWPT 273

Qy 181 PTLKGNVKSCANDMRPRITPDDLEDQPATNKTFDVKHSETINISLDTLDCSRIDGQPHT 240  
 |||  
 Db 274 PTLKGNVKSCANDMRPRITPDDLEDQPATNKTFDVKHSETINISLDTLDCSRIDGQPHT 333

Qy 241 PLNKTTHIVHATHTRALACIHEEEGSPPRPTPTKSAIHD LKRD IKLVGSPLRKYSESMKD 300  
 |||  
 Db 334 PLNKTTHIVHATHTRALACIHEEEGSPPRPTPTKSAIHD LKRD IKLVGSPLRKYSESMKD 393

Qy 301 LSLSPQTKYAIQGSMPNLNEMKIRSIEQNRYYQEQQIQIKAKDLNSSSSSEASLAGQQE 360  
 |||  
 Db 394 LSLSPQTKYAIQGSMPNLNEMKIRSIEQNRYYQEQQIQIKAKDLNSSSSSEASLAGQQE 453

Qy 361 FLFNHSEILAQSSRFNLHEVGRKSVKGS PVKNPHKRRSHELFS DAPS NESLYRNETVAI 420  
 |||  
 Db 454 FLFNHSEILAQSSRFNLHEVGRKSVKGS PVKNPHKRRSHELFS DAPS NESLYRNETVAI 513

Qy 421 SPPKKQ RVEDTTLP RSAAPANASARSSSAHAWPHAQSKKF KLAQTMSLMKKPATPRKVRD 480  
 |||  
 Db 514 SPPKKQ RVEDTTLP RSAAPANASARSSSAHAWPHAQSKKF KLAQTMSLMKKPATPRKVRD 573

Qy 481 TSIQPSVKLYDSELYMQTCINPDPFAATTTIDPFLASTMYLDEQAVDRHQADF KKWLNAL 540  
 |||  
 Db 574 TSIQPSVKLYDSELYMQTCINPDPFAATTTIDPFLASTMYLDEQAVDRHQADF KKWLNAL 633

Qy 541 VSIPADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNYLT KYRLET LRKA AVELFF 600  
 |||  
 Db 634 VSIPADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNYLT KYRLET LRKA AVELFF 693

Qy 601 SEQMRLPCSKVAVYV NKQALRIRSDRN LHL DVMQRTILELLLCFNPLWLRLGLEVV FGE 660  
 |||  
 Db 694 SEQMRLPCSKVAVYV NKQALRIRSDRN LHL DVMQRTILELLLCFNPLWLRLGLEVV FGE 753

Qy 661 KIQMQSNRDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYAETIKKHS LQKILFLLPFL 720  
 |||  
 Db 754 KIQMQSNRDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYAETIKKHS LQKILFLLPFL 813

Qy 721 DQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGDITRELRLRGYVLQHRQTF 780  
 |||  
 Db 814 DQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGDITRELRLRGYVLQHRQTF 873

Qy 781 LDEFDYAFNNLAVDLRDGVRLTRVVEVILLRDDLTRQLRVP AISRLQRI FNVKLALGALG 840  
 |||: |||  
 Db 874 LDEFDYAFNNLAVDLRDGVRLTRVMEVILLRDDLTRQLRVP AISRLQRI FNVKLALGALG 933

Qy 841 EANFQLGGDIAAQDIVDGHREKTL SLLWQLIYKFRSPKFHAAATVLQKWWRRHWLHVVIQ 900  
 |||  
 Db 934 EANFQLGGDIAAQDIVDGHREKTL SLLWQLIYKFRSPKFHAAATVLQKWWRRHWLHVVIQ 993

Qy 901 RRIRHKELMRRHRAATVIQAVFRGHQMRKYVKLFKTERTQAAILQKFTRRYLAQKQLYQ 960

Db	994	RRIRHKELMRRHRAATVIQAVFRGHQMRKYVKLFKTERTOAAIILQKFTRRYLAQKQLYQ	1053
Qy	961	SYHSIITIQRWWRAQQQLGRQHRQRFVELREAAIFLQRIWRRRLFAKKLLAAAETARLQRS	1020
Db	1054	SYHSIITIQRWWRAQQQLGRQHRQRFVELREAAIFLQRIWRRRLFAKKLLAAAETARLQRS	1113
Qy	1021	QKQQAASYSIQMQWRTYQLGRIQRHEFLRQDLMFVQRRMRSKWSMLEQRKEFQQLKRA	1080
Db	1114	QKQQAASYSIQMQWRSYQLGRIQRQQFLRQDLMFVQRRMRSKWSMLEQRKEFQQLKRA	1173
Qy	1081	AINIQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATIOMRIDRNHYYSLRKNVICLQQ	1140
Db	1174	AINIQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATIOMRIDRNHYYSLRKNVICLQQ	1233
Qy	1141	RLRAIMKMREQRENYLRLRNASILVQKRYMRQOMIQDRNAYLRTRKCIINVQRRWRATL	1200
Db	1234	RLRAIMKMREQRENYLRLRNASILVQKRYMRQOMIQDRNAYLRTRKCIINVQRRWRATL	1293
Qy	1201	QMRREKKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKER	1260
Db	1294	QMRREKKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQLKKVTLVVQKRRRALLQMRKER	1353
Qy	1261	QEYLHLREVTIKLQRRFHAQKSMRFMRAKYRGTOAAVSCLOMHWNRHLLRKRENSFLQL	1320
Db	1354	QEYLHLREVTIKLQRRFHAQKSMRFMRAKYRGTOAAVSCLOMHWNRHLLRKRENSFLQL	1413
Qy	1321	RQAAITLQRRYRARNMIKQLKSYAQLKQAAITIQTRYRAKKAMQKQVVLVYQKQREAIK	1380
Db	1414	RQAAITLQRRYRARNMIKQLKSYAQLKQAAITIQTRYRAKKAMQKQVVLVYQKQREAIK	1473
Qy	1381	VQRRYRGNLEMRKQIEVYQKQRAVIRLQKWWRSIRDMRLCKAGYRRIRLSSLSIQRKWR	1440
Db	1474	VQRRYRGNLEMRKQIEVYQKQRAVIRLQKWWRSIRDMRLCKAGYRRIRLSSLSIQRKWR	1533
Qy	1441	ATVQARRQREIFLSTIRKVRMLQAFIRATLLMRQQRREFEMKRRAAVVIQRRFRARCAML	1500
Db	1534	ATVQARRQREIFLSTIRKVRMLQAFIRATLLMRQQRREFEMKRRAAVVIQRRFRARCAML	1593
Qy	1501	KARQDYQLIQSSVILVQRKFRANRSMQARQEFVQLRTIAVHLQKFRGKRLMIEQRNCF	1560
Db	1594	KARQDYQLIQSSVILVQRKFRANRSMQARQEFVQLRTIAVHLQKFRGKRLMIEQRNCF	1653
Qy	1561	QLLRCSMPGFQARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQ	1620
Db	1654	QLLRCSMPGFQARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQ	1713
Qy	1621	GLLDIRKRIAQLRQEAKAVNSVRCKVQEA VRFLRGRFIASDALAVLSQLDRLSRTVPHLL	1680
Db	1714	GLLDIRKRIAQLRQEAKAVNSVRCKVQEA VRFLRGRFIASDALAVLSRLDRLSRTVPHLL	1773
Qy	1681	MWCSEFMSTFCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQM	1740
Db	1774	MWCSEFMSTFCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQM	1833
Qy	1741	LLRWCDKDSEIFNTLCTLIWVFAHCPKKRKIIHDYMTNPEAIYMVRETKKLVARKEKMKQ	1800

## RESULT 2

```

ID      Q7QAG9          PRELIMINARY;          PRT; 1538 AA.
AC      Q7QAG9;
DT      01-MAR-2004 (TrEMBLrel. 26, Created)
DT      01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      AgCP7874 (Fragment).
GN      Name=agCG45985; ORFNames=ENSANGG00000018773;
OS      Anopheles gambiae str. PEST.
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX      NCBI_TaxID=180454;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=PEST;
RA      Anopheles Genome Sequencing Consortium;
RL      Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC      -!- CAUTION: The sequence shown here is derived from an
CC      EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC      preliminary data.
DR      EMBL; AAAB01008888; EAA08939.1; -.
DR      InterPro; IPR001715; Calponin-like.
DR      InterPro; IPR000048; IQ_region.
DR      Pfam; PF00307; CH; 1.
DR      Pfam; PF00612; IQ; 31.
DR      PROSITE; PS50021; CH; 1.
DR      PROSITE; PS50096; IQ; 6.
FT      NON_TER      1      1
FT      NON_TER      1538      1538
SQ      SEQUENCE      1538 AA; 183217 MW; 0604DF326B743CC1 CRC64;

```

Qy	463	AQTMSLMKKPATPRKVRDTSIQPSVKLYDSELYMQTCINPDFAATTTTIDPFLASTMYLD	522
		:::    :     :   :         : ::                :	
Db	3	SRSLSL-KRTAVPCSLPPKSEEKRVFLYDSDRHLKTLINPDFAATTCNPFLTVMYLD	61
Qy	523	EQAVDRHQADFKKWLNALVSI PADLDADLNNKIDVGKLFNEVRNKELVVAPTKEEQSMNY	582
		:  ::::                     :   :         :     :	
Db	62	ERAFEQYERQMKKWLNALVTIPADLDTEPNKPLDVGKLFDEVKSKELTLAPTKELISSKY	121
Qy	583	LTKYRLETLRKA AVELFFSEQMRLPCKSVAVYVVKQALRIRSDRNHLHLDVVMQRTILELL	642
		:   :     : :       :     :           :   :	

Db	122	Y-KTRLNHLRSAGIALYTSEEIAMPLRKVAAQIEKQLLSRLTRDRLHLDLVLQRSILELL	180
Qy	643	LCFNPLWLRLGLEVVFGKEIQMQSNRNDIVGLSTFILNRLFRNKCEEQRYSKAYTLTEEYA	702
Db	181	LCFNPLWLRLGLEVVFGQIELQSNRNDIVGLSTFIIHRLFRDRYLEARNSKAYNLSRAYA	240
Qy	703	ETIKKHSLOKILFLLPFLDQAKQKRIVKHNPCLFVKKSPHKETKDILLRFSSELLANIGD	762
Db	241	EHMRKFTLRMVLFLLFLDQAKRKRKLIKHNPCLFVRNAPHKETKEILIRFASQLVSGIGD	300
Qy	763	ITRELRLRGYVLQHRQTFLDEFDYAFNNLAVDLRDGVRLTRVVEVILLRDDLTRQLRVPA	822
Db	301	ITKHMKRVGIVLSHKQSFLEDEYNYAFENLAVDLRDGVRLTRVMEIILLRDDLRSASLRVPP	360
Qy	823	ISRLQRIQFNVKLALGALGEANFQLGGDIAAQDIVDGHREKTLSSLWQLIYKFRSPKFHAA	882
Db	361	ISRLQKIHNINLALVALEQADYKIAGNVTAKDIDGHREQTMSLLWQIVYKFRAPKFNA	420
Qy	883	ATVLQKWWRRLHVLHVVIQRRIRHKELMRHRRAATVIQAVFRGHQMRKYVKLFKTERTQAA	942
Db	421	AIVLQRWWRMNWLKVTISRRIEEKRALRREAAARTIQAAVRGYCVRVWYEAHRRQKLR	480
Qy	943	IILQKFTRRYLAQKQLYQSYHSIITIQRWWRAQQLGRQHRQRFVELREAAIFLQRIWRR	1002
Db	481	VTIQRFSRRYLAQKLAARRFSAIVRIQQWWRVTRQMRQARERFLCRKSAIVLQTSYRR	540
Qy	1003	LEAKKLLAAAE-----TARLQRSQKQ	1023
Db	541	ALGRKLLAAATLIGQIRAEAKHRHLQATIIQRSIKSYVIHRRHLATVNGMVAFIRRKRLQ	600
Qy	1024	QAAASYIQMQWRTYQLGRIQRHEFLRQDRLIMFVQRRMRSKWSMLEQRKEFQQLKRAAIN	1083
Db	601	NRSAAKIQ---AYQRMRIVRKEYLRSRSAACIQRRWRECMEARQLRNRFLMRASAIR	656
Qy	1084	IQQRWRAKLSMRKCNADYLALRSSVLKVQAYRKATIQRIDRNHYSLRKNVICLQQRRL	1143
Db	657	LQQQYRGWRQMRQDRHTYANARNLIVQVQRRWRGTLAMRKERANYRTLRRVTINVQRRFR	716
Qy	1144	AIMKMREQRENYLRLRNASILVQKRYRMRQMIQDRNAYLRTRKCIINVQRRWRATLQMR	1203
Db	717	ARQAMQSEVERYRTLCKATVTLQQRFRANKAMMEQRQYNSLRVATLCVQRRFRAQLSMR	776
Qy	1204	REKKNYLHLQTTTKRIQIKFRAKREMKKQRAEFLQKKVTLVVQKRRRALLQMRKERQEY	1263
Db	777	AARASYAKVRCAILTIQSQYRATLAMRHARDRFVTLRRCITITVQARFRAILAGRAAKQRY	836
Qy	1264	LHLREVTIKLQRRFHAQKSMRFMRKYRGTAQAVSCLQMHWRNHLLRKRERNSEFL-----	1318
Db	837	ESIRKATLHIQKWRATLEMRQVRSYHRQCNAAALTQRSWGRVLLQKRFHHDYLLYRGA	896
Qy	1319	-----QLRQAAITLQRRYRARNMIKQLKSYAQLKQAAITIQT	1356
Db	897	ATVLQRRYRALVQGRMVREMQHCRWAAVTIQRRRLRATLQMNDRKAFQLRQSVLVVQR	956
Qy	1357	RYRAKKAMQKQVVLV-----QKQR-----EAIKQVRRYRGNL	1389
Db	957	RFRANRACRVQRVQYAALKRSAITISHRWAATLHMRQQRSDFLRLKSATVVMQRRYRAQR	1016



Qy 1390 EMRKQIEVYQKQKQAVIRLQKWWRSIRDMLCKAGYRRIRLSSLSIQRKWRATVQARRQR 1449  
 :: :: |:: | |:: ||: :: | | : : :: :: | : | |

Db 1017 AKQQAVQQYERMRAAIVLLQKRYAQRAMEKCRGRFLNLKSASIVVQEFYRGYRNMHRDR 1076

Qy 1450 EIFLSTIRKVRMLQAFIRATLLMRQQRREFEMKRRAAVVIQRRFRARCAMLKARQDYQLI 1509  
 | : | : | | | | : | | : | : | | | | | : : :

Db 1077 AAFIRLRESVLAIQRRFRGKLLTRQTVVEYERKRKAATLQRWFRGHRAMVKERKRFTAL 1136

Qy 1510 QSSVILVQRKFRANR-----SMKQARQEFVQ----LRT 1538  
 : : :: | :: | | | | | | | | : | : | : |

Db 1137 KHAATVIQARYRARRAMLHDAAQYGRIYRAVTTIQIHWKATLSMRERDRYQOTLRAVRT 1196

Qy 1539 IAVH-----LQQKFRGKRLMIEQRNC----FQLLRCSMPGFQ 1571  
 | : | : | : | : | : | : | : | : | : | : |

Db 1197 IQIHYRAYRERLIDEANYRIYRSIVVQRRYREKL----QTRCESLRFEQIRRTVRGLQ 1252

Qy 1572 ARARGFMARKRFQALMTPEMMDLIRQKRAAKVIQRYWRGYLIRRRQKHQGLLDIRKRIAQ 1631  
 | | : | : | | | : | : | : | | | : | : | : | : |

Db 1253 TYGRGVLARRAFLALLTPEYLERKQKQKAAALRIQAWWRGAYHRKRYQTMQMRKIAQQMVA 1312

Qy 1632 LRQEAKAVNSVRCK--VQEAVRFLRGRFIASDALAVLSQLDRLSRTVPHLLMWCSEFMST 1689  
 | | : : | : : | : | : | : | : | : | : | : |

Db 1313 SRMAARRDPTIRLSNVSRCLRLKTRFSSEAGILKRLEMSRLVPHLLMEDAVFLSV 1372

Qy 1690 FCYGIMAQAIRSEVDKQLIERCSRIILNLARYNSTTVNTFQEGGLVTIAQMLLRWCDKDS 1749  
 | | | : | | | | | | | | | | | : | | | | | : | | | | |

Db 1373 FCYNMMAQAIRSEVDKILIEICARIILNLARFRGTKEQAFQEDGLVTVSQMLLRWCDKDC 1432

Qy 1750 EIFNTLCTLIWVFAHCPKRRKIIHDYMTNPEAIYMVRETKKLVARKEKMKQNARKP----- 1805  
 | | : | | | : | | | : : | | : | | | | | | | : | : |

Db 1433 GIFSTLCTLLWVLAHDNKKNAIRRYMISKDAIYMLRETKKLVQRKEKMRKNVQRPVGCL 1492

Qy 1806 -----PPMTSGRYKSQKINFPCSLPSLEPDFGIIRYSPYTFISSVYAFDTILCKLQIDM 1860  
 | : : | | | | | : | | | | | : | : | : | : |

Db 1493 VAPNPQLMR-----TVPSLEPDFGVNRSKPYVFYSSVFGFERVLQMLEVDL 1538